External Validation and Comparison of Prostate Cancer Risk Calculators Incorporating Multiparametric Magnetic Resonance Imaging for Prediction of Clinically Significant Prostate Cancer

Saba, Karim ; Wettstein, Marian S ; Lieger, Laura ; Hötker, Andreas M ; Donati, Olivio F ; Moch, Holger ; Ankerst, Donna P ; Poyet, Cédric ; Sulser, Tullio ; Eberli, Daniel ; Mortezavi, Ashkan

Abstract: PURPOSE: To externally validate recently published prostate cancer risk calculators (PCa-RCs) incorporating multiparametric magnetic resonance imaging (mpMRI) for the prediction of clinically significant prostate cancer (csPCa) and compare their performance to mpMRI-naïve PCa-RCs. MATERIAL AND METHODS: Men without previous PCa diagnosis undergoing transperineal template saturation prostate biopsy with fusion-guided targeted biopsy between 11/2014 and 03/2018 in our academic tertiary referral center were identified. Any Gleason pattern 4 was defined to be csPCa. Predictors (age, PSA, DRE, prostate volume, family history, previous prostate biopsy and highest region of interest according to PIRADS) were retrospectively collected. Four mpMRI-PCa-RCs and two mpMRI-naïve PCa-RCs were evaluated for their discrimination, calibration and clinical net benefit using a ROC analysis, calibration plots and a decision curve analysis, respectively. RESULTS: Out of 468 men, 193 (41%) were diagnosed with csPCa. Three mpMRI-PCa-RCs showed similar discrimination with area-underneath-the-receiver-operating-characteristic-curves (AUC) from 0.83 to 0.85, which was significantly higher than the other PCa-RCs (AUCs: 0.69-0.74). Calibration-in-the-large showed minimal deviation from the true amount of csPCa by 2% for two mpMRI-PCa-RCs, while the other PCa-RCs showed worse calibration (11-27%). A clinical net benefit could only be observed for three mpMRI-PCa-RCs at biopsy thresholds 15%, while none of the six investigated PCa-RCs demonstrated clinical utility against a biopsy all strategy at thresholds <15%. CONCLUSIONS: Performance of the mpMRI-PCa-RCs varies, but they generally outperform mpMRI-naïve PCa-RCs in regard to discrimination, calibration and clinical usefulness. External validation in other biopsy settings is highly encouraged.

DOI: https://doi.org/10.1097/JU.0000000000000622

External Validation and Comparison of Prostate Cancer Risk Calculators Incorporating Multiparametric Magnetic Resonance Imaging for Prediction of Clinically Significant Prostate Cancer

Saba K, Wettstein MS, Lieger L, Hötker AM, Donati OF, Moch H, Ankerst DP, Poyet C, Sulser T, Eberli D, and Mortezavi A

DOI: 10.1097/JU.0000000000000622
Reference: JU-19-1515

To appear in: *The Journal of Urology*
Accepted Date: 17 October 2019

Please cite this article as: Wang Saba K, Wettstein MS, Lieger L, Hötker AM, Donati OF, Moch H, Ankerst DP, Poyet C, Sulser T, Eberli D, and Mortezavi A, External Validation and Comparison of Prostate Cancer Risk Calculators Incorporating Multiparametric Magnetic Resonance Imaging for Prediction of Clinically Significant Prostate Cancer, *The Journal of Urology®* (2018), doi: 10.1097/JU.0000000000000622.

The Journal of Urology Publish Ahead of Print
DOI: 10.1097/JU.0000000000000622

External Validation and Comparison of Prostate Cancer Risk Calculators Incorporating Multiparametric Magnetic Resonance Imaging for Prediction of Clinically Significant Prostate Cancer

Karim Saba[a *], Marian S. Wettstein[a,b *], Laura Lieger[a], Andreas M. Hötker[c], Olivio F. Donati[c], Holger Moch[d], Donna P. Ankerst[e], Cédric Poyet[a], Tullio Sulser[a], Daniel Eberli[a], Ashkan Mortezavi[a]

[a] Department of Urology, University Hospital Zurich, University of Zurich, Zurich, Switzerland
[b] Institute for Health Policy, Management and Evaluation, University of Toronto
[c] Institute of Diagnostic and Interventional Radiology, University Hospital Zurich, University of Zurich, Zurich, Switzerland
[d] Department of Pathology and Molecular Pathology, University Hospital Zurich, University of Zurich , Zurich , Switzerland
[e] Department of Mathematics, Technical University of Munich, Garching, Munich, Germany

karim.saba@usz.ch
marian.wettstein@me.com
laura.lieger@usz.ch
andreas.hoetker@usz.ch
olivio.donati@usz.ch
holger.moch@usz.ch
ankerst@tum.de
cedric.poyet@usz.ch
tullio.sulser@usz.ch
daniel.eberli@usz.ch
ashkan.mortezavi@usz.ch

Karim Saba and Marian S. Wettstein contributed equally to this work.

Corresponding author:

PD Dr. med. Ashkan Mortezavi
Department of Urology
University Hospital Zürich
Frauenklinikstrasse 10
8091 Zürich
Tel. +41 43 253 88 27
Fax. +41 44 255 45 55
ashkan.mortezavi@usz.ch

Key words (MeSH): Prostate Cancer, Magnetic Resonance Imaging, Multivariable Risk Stratification, Risk Calculator, Biopsy

Abstract

Purpose: To externally validate recently published prostate cancer risk calculators (PCa-RCs) incorporating multiparametric magnetic resonance imaging (mpMRI) for the prediction of clinically significant prostate cancer (csPCa) and compare their performance to mpMRI-naïve PCa-RCs.

Material and Methods: Men without previous PCa diagnosis undergoing transperineal template saturation prostate biopsy with fusion-guided targeted biopsy between 11/2014 and 03/2018 in our academic tertiary referral center were identified. Any Gleason pattern ≥4 was defined to be csPCa. Predictors (age, PSA, DRE, prostate volume, family history, previous prostate biopsy and highest region of interest according to PIRADS) were retrospectively collected. Four mpMRI-PCa-RCs and two mpMRI-naïve PCa-RCs were evaluated for their discrimination, calibration and clinical net benefit using a ROC analysis, calibration plots and a decision curve analysis, respectively.

Results: Out of 468 men, 193 (41%) were diagnosed with csPCa. Three mpMRI-PCa-RCs showed similar discrimination with area-underneath-the-receiver-operating-characteristic-curves (AUC) from 0.83 to 0.85, which was significantly higher than the other PCa-RCs (AUCs: 0.69-0.74). Calibration-in-the-large showed minimal deviation from the true amount of csPCa by 2% for two mpMRI-PCa-RCs, while the other PCa-RCs showed worse calibration (11-27%). A clinical net benefit could only be observed for three mpMRI-PCa-RCs at biopsy thresholds ≥15%, while none of the six investigated PCa-RCs demonstrated clinical utility against a *biopsy all* strategy at thresholds <15%.

Conclusions: Performance of the mpMRI-PCa-RCs varies, but they generally outperform mpMRI-naïve PCa-RCs in regard to discrimination, calibration and clinical usefulness. External validation in other biopsy settings is highly encouraged.

Introduction

Prostate-specific antigen (PSA) screening leads to more diagnoses of localized instead of advanced prostate cancer (PCa)[1], but has also resulted in overdiagnosis of clinically insignificant PCa[2]. Instead of relying on PSA and digital rectal examination (DRE) as the sole criteria to biopsy men, use of multivariable PCa risk calculators (PCa-RCs) provide possibly more accurate predictions for PCa[3] and may hereby reduce the number of negative prostate biopsies and overdiagnosis[4]. Multiparametric magnetic resonance imaging (mpMRI) has been shown to reduce the number of unnecessary prostate biopsies, when used as a triage test[5]. Furthermore, mpMRI has recently been incorporated into PCa-RCs to enhance their predictive ability for clinically significant prostate cancer (csPCa)[6–9]. It has been shown that certain mpMRI-naïve PCa-RCs perform worse in external validation studies, than would have been anticipated from their original reports[10,11]. Hence, we externally validated four novel PCa-RCs incorporating mpMRI[6–9] and compared their performances to two established mpMRI-naïve RCs[12,13].

Material and Methods

Study design and setting

All men who underwent mpMRI and transperineal template saturation prostate biopsy with additional fusion-guided targeted biopsy for suspicion of csPCa from 11/2014–03/2018 in an academic tertiary referral center were considered for this study. Transrectal ultrasound–mpMRI fusion and virtual needle placement were done using the BiopSEE® software (MedCom, Darmstadt, Germany) for systematic biopsy as reported previously[14]. Additional fusion-guided targeted biopsies (2-4 cores) were performed for each region of interest (ROI) classified as grade ≥3 according to the Prostate Imaging Reporting and Data System (PIRADS) version 2. A specialized uro-pathologist assessed each core separately. Any Gleason pattern ≥4 was defined to be csPCa.
Patients either underwent mpMRI at our institution (including triplanar T2- and diffusion-weighted and dynamic contrast-enhanced sequences) or were referred to our institution after having mpMRI performed externally. In case Likert scale or PIRADSv1 was used, internal reassessment according to PIRADSv2 guidelines[15] was done by specialized uro-radiologists. Cases with qualitative insufficiently performed MRIs were excluded. Predictors (age, PSA, DRE, prostate volume, family history, previous biopsy and highest ROI-grade on mpMRI) were retrospectively retrieved from electronic medical records. The study was approved by the local ethics committee.

Analysis

We investigated four PCa-RCs incorporating mpMRI (Radtke et al.[9], van Leeuwen et al.[6], Mehralivand et al.[7] and mpMRI-ERSPC-RC[8]) and two conventional mpMRI-naïve PCa-RCs (ERSPC-RC[12] and PBCG-RC[13]). Specifics of the corresponding studies are outlined in the Supplementary Table. For three PCa-RCs[6,7,9] predictions were calculated based on published logistic regression models. For the other PCa-RCs corresponding authors, who were blinded to biopsy results, provided the full regression model[13] or calculated predictions[8,12]. Although family history information was missing for almost half of our cohort, predictions could still be calculated for all patients as this predictor is only inquired by the PBCG-RC as an optional input variable. For other marginally missing values such as highest grade according to PIRADS and DRE no missingness assumptions were made (complete case analysis).

We performed two sensitivity analyses: First, we investigated the effect of missing family history information on the performance of the PBCG-RC by repeating the analysis among patients with complete and among an imputed dataset. Second, the performance of the PCa-RC of Mehralivand et al. was reanalyzed in a cohort restricted to cases with positive mpMRI (PIRADS grade ≥3) to account for the peculiar low amount of PIRADS grade <3 (8.1%) of its development cohort. In addition, we evaluated potential performance improvements by recalibration as described by Strobl and colleagues[16,17]. For this, two thirds of our cohort (training cohort) was used to update the models while the remaining third was used as a validation cohort.

For each PCa-RC we assessed discrimination, calibration and clinical usefulness. Discrimination was evaluated by area-under-the-receiver-operating-characteristics-curves (AUC). Calibration-in-the-large was assessed by comparing the predicted proportion of csPCa from each PCa-RC to the proportion observed in our cohort. We further investigated model calibration both numerically (intercept and slope) and visually (calibration plot). Clinical usefulness was assessed by a decision curve analysis (DCA)[18]. All analyses were performed in R 3.4.4 (R Foundation for Statistical Computing, Vienna, Austria) using the *mice* (multiple imputation)[19], *pROC* (AUC)[20] and *rms* (Calibration plots)[21] packages. The DCA was implemented by the published code of Vickers et al.[18]

Results

A total of 935 men were considered for this study. After excluding 401 patients with previously diagnosed PCa, 10 men with qualitatively insufficient MRIs for PIRADSv2 assessment and 56

individuals with missing DRE information, 468 men were left for analysis. Table 1 summarizes the cohorts' clinicopathological characteristics and biopsy results. In our cohort, median age was 64.5 years (interquartile range [IQR]: 59-68.9), median PSA was 6.6 ng/ml (IQR: 4.5-9.9) and median prostate volume was 48.9ml (IQR: 36-63.7). There were 145 (31%) men who had at least one previous negative prostate biopsy. In 87 (18.5%) cases no ROI was seen on mpMRI. Highest ROI was graded to be PIRADS 2, 3, 4 and 5 in 68 (14.5%), 94 (20%), 162 (34.6%) and 57 (12.2%) cases, respectively. Median systematic biopsy core number was 40 (IQR: 39-42) and median number of targeted biopsy cores per ROI was 3 (IQR: 2-3). Gleason Score 3+3=6 PCa was found in 55 (11.8%) patients and csPCa was diagnosed in 193 (41.2%) Patients.

AUCs with corresponding 95%-confidence intervals (95%-CI) for PCa-RCs under evaluation are shown in Figure 1. The AUCs of three mpMRI-PCa-RCs were comparable (mpMRI-ERSPC-RC: AUC of 0.85 [95%-CI: 0.82-0.89]; Mehralivand et al.: 0.84 [95%-CI: 0.80-0.87] and van Leeuwen et al.: 0.83 [95%-CI: 0.80-0.87]), but were considerably higher compared to the mpMRI-PCa-RC of Radtke et al. (AUC of 0.73 [95%-CI: 0.69-0.78]) and the mpMRI-naïve PCa-RCs (ERSPC-RC: AUC of 0.74 [95%-CI: 0.69-0.79]; PBCG-RC: AUC of 0.69 [95%-CI: 0.65-0.74]).

Calibration-in-the-large and calibration plots for each PCa-RC (including intercept and slope indicating miscalibration and overfitting, respectively) are visualized in Figures 2 and 3. Calibration-in-the-large of the mpMRI-PCa-RCs of Radtke et al. (+2%) and van Leeuwen et al. (-2%) showed minimal deviation from the observed proportion of 41% csPCa in our cohort, while the mpMRI-ERSPC-RC (deviation of -14%) and the PBCG-RC (deviation of -11%) exhibited intermediate miscalibration. The mpMRI-PCa-RC of Mehralivand et al. and the mpMRI-naïve ERSPC-RC yielded large deviations of +27% and -29%, respectively. When it comes to overfitting, the mpMRI-PCa-RCs of Radkte et al. and van Leeuwen et al. produce too extreme predictions as reflected by their calibration slopes considerably smaller than 1 in comparison to the other PCa-RCs.

Clinical usefulness was evaluated using DCA (Figure 4). In a scenario where missing 10% of csPCa is acceptable, application of any PCa-RCs under evaluation even showed to be clinically harmful compared to the strategy to biopsy all. For a more commonly used clinical threshold probability of 15% – at which the downside of 5.6 unnecessary prostate biopsies is equated to missing one csPCa – use of three mpMRI-PCa-RCs (mpMRI-ERSPC-RC, the one from van Leeuwen et al. and from Mehralivand et al.) exhibited a clinical net benefit. At this threshold the use of the mpMRI-ERSPC-RC and the PCa-RC by van Leeuwen et al. would omit 84 and 67 per 1000 prostate biopsies, respectively, without missing one case of csPCa. Use of the other three PCa-RCs under evaluation exhibited net harm in this scenario.

To explore whether the worse performance of the mpMRI-naïve PBCG-RC is potentially driven by the high proportion of missing family history information, we performed a sensitivity analysis involving only patients with available family history (n=278) and a multiple imputed dataset (n=468). Neither the discrimination ($AUC_{\text{available family history}}$: 0.67, AUC_{imputed}: 0.70) nor the calibration ($\text{intercept}_{\text{available family history}}$: 0.35, $\text{intercept}_{\text{imputed}}$: 0.40, $\text{slope}_{\text{non-missing family history}}$: 0.70, $\text{slope}_{\text{imputed}}$: 0.78) were substantially improved. We further hypothesized that the mpMRI-PCa-RC of Mehralivand et al. would perform better for patients with a PIRADS grade ≥3, in accordance with its development cohort characteristics (92% with PIRADS grade ≥3). Our sensitivity analysis with mpMRI positive patients only (PIRADS ≥3; n=313) showed no improvement of discrimination (AUC of 0.81 [95%-CI: 0.76-0.85]), but an improvement in calibration (intercept: 1.26). Our split-sample recalibration/validation approach demonstrated considerable improvement of predictive performances for most PCa-RCs (Supplementary Figures 1-4).

Discussion

In this external validation study we comprehensively assessed all currently available mpMRI-PCa-RCs and compared them to two established mpMRI-naïve PCa-RCs. In general, the PCa-RCs incorporating mpMRI outperformed their mpMRI-naïve predecessors. Among all risk models under evaluation the mpMRI-PCa-RCs of van Leeuwen et al. and the mpMRI-ERSPC-RC showed a distinct clinical net benefit at a threshold of 15%.

It has been shown that use of PCa-RCs can reduce overdiagnosis of clinically insignificant PCa at a small expense of missing csPCa compared to using a certain PSA-threshold for the decision to biopsy men. The use of such PCa-RCs in daily routine can be facilitated greatly by their online dissemination – as has been done for the ones based on the PSA-screening trials PCPT and ERSPC[12,22,23]. The PCa-RC of the Prostate Biopsy Collaborative Group (PBCG) is based on opportunistic screening cohorts and is propagated online as the *quasi* successor to the PCPT-RC[13]. Validation was performed with three large mpMRI-naïve European cohorts (n=10'377), but the PBCG-RC has not been independently validated so far. With performance of mpMRI becoming the *de facto* standard for patients at risk in many urological care centers and recent EAU guidelines approval[24], radiological risk-assessment of csPCa is becoming increasingly available before prostate biopsy. Novel PCa-RCs incorporate mpMRI as an additional parameter, as the consequential next step to enhance their performance. Results from the PRECISION trial attest to the validity of mpMRI as a triage test for targeted biopsies in biopsy-naïve patients[5]. However, it does not address the risk of csPCa in men with negative mpMRI. The median negative predictive value (NPV) of mpMRI is 80.4 and 88.2% for biopsy-naïve patients and in a repeat biopsy setting, respectively, according to a recent meta-analysis[25], but is strongly dependent on prevalence and therefore not generalizable. The current EAU guidelines put emphasis on clinical risk-assessment and shared decision making for how to proceed further with patients with negative mpMRI[24]. Wang and colleagues found individual risk assessment with PCPT-RC to correlate well with risk of csPCa in patients with negative mpMRI[26], highlighting the potential role of PCa-RCs to select patients with low NPV of mpMRI who should receive biopsy notwithstanding[24]. If mpMRI is positive, individual risk assessment can strengthen the shared decision making process by *conferring agency* in well-informed patients[27]. Original reports of these mpMRI-PCa-RCs show encouraging discriminative performances (AUCs: 0.81-0.88)[6–9] compared to the original reports of the mpMRI naïve ERSPC-RC and PBCG-RC (AUCs: 0.71-0.81)[12,13]. However, all risk models are dependent on multiple factors peculiar to their development cohorts, i.e. disease prevalence and biopsy method, hindering their direct generalizability. In line with this, validation studies of mpMRI-naïve PCa-RCs reported less optimistic results than expected from their original reports[10,28]. Of note, the mpMRI-PCa-RCs of Mehralivand et al. and van Leeuwen et al. included an external validation in their original reports. Our study represents the first independent external validation study of mpMRI-PCa-RCs and their comparison to two established mpMRI-naïve PCa-RCs. In this effort we investigated all currently available mpMRI-PCa-RCs for csPCa to the best of our knowledge.

The better discrimination of the PCa-RCs incorporating mpMRI can be simply explained by the additional information generated by the inherent diagnostic accuracy of the mpMRI[29]. The lower discriminative performance of the PBCG-RC (AUC: 0.69) might be further explained by its omission of prostate volume as a predictor. This hypothesis is in line with Ankerst and colleagues showing improved discrimination of the PCPT-RC by adding prostate volume as an additional predictor[30]. The hypothesis that the low discriminative performance of the PBCG-RC is caused by the high proportion of missing information on family history could not be confirmed in a sensitivity analysis. Differences of ethnical background between our cohort and some of the PBCG-RC' development cohorts (i.e. the Durham cohort consisting of >60% men with African American ancestry) is another possible explanation for the underperformance[13].

The underprediction of the ERSPC-RC, PBCG-RC and mpMRI-ERSPC-RC is mainly driven by the lower prevalence of csPCa in their development cohorts of 4.5-35%. Among all risk models, the PCa-RC of Mehralivand et al. showed a prominent overprediction. Our hypothesis that the overestimation partly stems from the higher proportion of positive mpMRIs of 82.1% compared to 61.8% in our cohort was confirmed in a sensitivity analysis showing improved calibration after excluding patients with PIRDAS <3.

The minimal deviation in the calibration-in-the-large of the mpMRI-PCa-RCs from Radtke et al. and van Leeuwen et al. to our cohort could be explained by their systematic transperineal biopsy protocols with high coverage (median of 24 and 30 cores, respectively) similar to our approach (median of 40 systematic biopsy cores), compared to the 6-12 systematic cores biopsies used to the develop the other PCa-RCs.

Combining discrimination, calibration and risk aversion in a clinical utility perspective, only the mpMRI-ERSPC-RC and the PCa-RC of van Leeuwen et al. demonstrated a distinct net benefit in our validation setting when a risk of a false-negative prediction of 15% is accepted. Application of the remaining PCa-RCs were even harmful in comparison to a *biopsy all men* strategy or manifested a clinical net benefit only at clinically implausible risk thresholds.

The strength of our study is the comprehensive assessment (discrimination, calibration and clinical utility) of all currently available risk models for csPCa with mpMRI. Compared to other validation studies, we used a highly accurate gold standard for outcome ascertainment (transperineal saturation biopsy with a median of 40 systematic biopsy cores and fusion-guided targeted biopsy).

Our study has limitations. Foremost, it is a retrospective study. A selection bias due to already previously performed risk assessments has to be considered. Furthermore, family history assessment was missing in 40%, which could have led to inferior performance of the PBCG-RC, although a sensitivity analysis (complete cases and multiple imputation analysis) showed no improvement. Since all our patients have Caucasian descent, our findings are not fully generalizable to more heterogeneous populations.

Our results affirm the importance of choosing the *best* PCa-RC for a specific setting, which consequently warrants replication of our study in other centers with different screening, mpMRI and biopsy practices. In an idealistic setting, replication of our study in each Urologist's own patient cohort would lead to the optimal choice. Although online dissemination of PCa-RC greatly facilitates their use in daily practice, risk models with unpublished regression coefficients hinder further much needed validation studies for different biopsy settings. Although recalibration led to expected predictive improvements for most PCa-RCs, this approach is usually not available in clinical practice and is not within the scope of this investigation.

Conclusions

In our external validation setting the best performances in regard to discrimination, calibration and clinical utility were achieved by PC-RCs incorporating mpMRI. From a clinical utility perspective the mpMRI-PCa-RCs of van Leeuwen et al. and the mpMRI-ERSPC-RC outperformed the other models under evaluation. However, clinical net benefit could only be observed for non risk-averse men when a relatively high threshold range above 15% was applied. As clinical utility is strongly cohort dependent, but the use of certain PC-RCs facilitated greatly by its online dissemination, there is a strong need for more external validation for different biopsy settings.

Acknowledgments

Monique J. Roobol supported this validation study by providing the predictions for the ERSPC-PCa-RC and mpMRI-ERSPC-PCa-RC.

Funding

MS.W. is funded by the Swiss Cancer Research Foundation (BIL KFS-4009-08-2016), the Professor Dr. Max Cloetta Foundation & Uniscientia Foundation (Vaduz), the Swiss Association of Urology and the Kurt & Senta Herrmann Foundation. A.M. receives funding from the Swiss Cancer League.

References

1. Ilic D, Neuberger M, Djulbegovic M, et al: Screening for prostate cancer (Review). 2013.

2. Draisma G, Etzioni R, Tsodikov A, et al: Lead time and overdiagnosis in prostate-specific antigen screening: Importance of methods and context. J. Natl. Cancer Inst. 2009; **101**: 374–383.

3. Schröder F and Kattan MW: The comparability of models for predicting the risk of a positive prostate biopsy with prostate-specific antigen alone: a systematic review. Eur. Urol. 2008; **54**: 274–90. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0302283808006258, accessed June 30, 2019.

4. Roobol MJ, Steyerberg EW, Kranse R, et al: A Risk-Based Strategy Improves Prostate-Specific Antigen-Driven Detection of Prostate Cancer. Eur. Urol. 2010; **57**: 79–85.

5. Kasivisvanathan V, Rannikko AS, Borghi M, et al: MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. N. Engl. J. Med. 2018; **378**: 1767–1777. Available at: http://www.nejm.org/doi/10.1056/NEJMoa1801993, accessed November 19, 2018.

6. van Leeuwen PJ, Hayen A, Thompson JE, et al: A multiparametric magnetic resonance imaging-based risk model to determine the risk of significant prostate cancer prior to biopsy. BJU Int. 2017; **120**: 774–781. Available at: http://www.ncbi.nlm.nih.gov/pubmed/28207981, accessed March 12, 2018.

7. Mehralivand S, Shih JH, Rais-Bahrami S, et al: A Magnetic Resonance Imaging–Based Prediction Model for Prostate Biopsy Risk Stratification. JAMA Oncol. 2018. Available at: http://www.ncbi.nlm.nih.gov/pubmed/29470570, accessed March 12, 2018.

8. Alberts AR, Roobol MJ, Verbeek JFM, et al: Prediction of High-grade Prostate Cancer Following Multiparametric Magnetic Resonance Imaging: Improving the Rotterdam European Randomized Study of Screening for Prostate Cancer Risk Calculators. Eur. Urol. 2018.

9. Radtke JP, Wiesenfarth M, Kesch C, et al: Combined Clinical Parameters and Multiparametric Magnetic Resonance Imaging for Advanced Risk Modeling of Prostate Cancer-Patient-tailored Risk Stratification Can Reduce Unnecessary Biopsies. Eur. Urol. 2017; **72**: 888–896. Available at: http://linkinghub.elsevier.com/retrieve/pii/S0302283817302671, accessed March 12, 2018.

10. Poyet C, Nieboer D, Bhindi B, et al: Prostate cancer risk prediction using the novel versions of the European Randomised Study for Screening of Prostate Cancer (ERSPC) and Prostate Cancer Prevention Trial (PCPT) risk calculators: independent validation and comparison in a contemporary E. 2016: 401–408.

11. Ankerst DP, Boeck A, Freedland SJ, et al: Evaluating the Prostate Cancer Prevention Trial High Grade prostate cancer risk calculator in 10 international biopsy cohorts: Results from the prostate biopsy collaborative group. World J. Urol. 2014; **32**: 185–191.

12. Roobol MJ, Van Vugt HA, Loeb S, et al: Prediction of prostate cancer risk: The role of prostate volume and digital rectal examination in the ERSPC risk calculators. Eur. Urol. 2012; **61**: 577–583.

13. Ankerst DP, Straubinger J, Selig K, et al: A Contemporary Prostate Biopsy Risk Calculator Based on Multiple Heterogeneous Cohorts. Eur. Urol. 2018; **74**: 197–203. Available at: https://doi.org/10.1016/j.eururo.2018.05.003.

14. Mortezavi A, Märzendorfer O, Donati OF, et al: Diagnostic Accuracy of Multiparametric Magnetic Resonance Imaging and Fusion Guided Targeted Biopsy Evaluated by Transperineal

Template Saturation Prostate Biopsy for the Detection and Characterization of Prostate Cancer. J. Urol. 2018; **200**: 309–318. Available at: http://www.ncbi.nlm.nih.gov/pubmed/29474846, accessed October 22, 2018.

15. Weinreb JC, Barentsz JO, Choyke PL, et al: PI-RADS Prostate Imaging – Reporting and Data System: 2015, Version 2. Eur. Urol. 2016; **69**: 16–40. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26427566, accessed December 26, 2018.

16. Strobl AN, Vickers AJ, Calster B Van, et al: Improving patient prostate cancer risk assessment□: Moving from static , globally-applied to dynamic , practice-specific risk calculators. 2015; **56**: 87–93.

17. Strobl AN, Thompson IM, Vickers AJ, et al: The Next Generation of Clinical Decision Making Tools: Development of a Real-Time Prediction Tool for Outcome of Prostate Biopsy in Response to a Continuously Evolving Prostate Cancer Landscape. J. Urol. 2015; **194**: 58–64. Available at: http://www.jurology.com/doi/10.1016/j.juro.2015.01.092, accessed October 8, 2019.

18. Vickers AJ and Elkin EB: Decision Curve Analysis□: A Novel Method for Evaluating Prediction Models. 2006: 565–574.

19. Buuren S van and Groothuis-Oudshoorn K: mice: Multivariate Imputation by Chained Equations in R. J. Stat. Softw. 2011; **45**: 1–67. Available at: http://www.jstatsoft.org/v45/i03/, accessed January 14, 2019.

20. Robin X, Turck N, Hainard A, et al: pROC: an open-source package for R and S+ to analyze and compare ROC curves.; 2011. Available at: http://expasy.org/tools/pROC/, accessed January 14, 2019.

21. Harrell F: Regression Modeling Strategies. Springer; 2016.

22. Roobol MJ, Steyerberg EW, Kranse R, et al: A Risk-Based Strategy Improves Prostate-Specific Antigen–Driven Detection of Prostate Cancer. Eur. Urol. 2010; **57**: 79–85. Available at: http://www.ncbi.nlm.nih.gov/pubmed/19733959, accessed December 30, 2018.

23. Ankerst DP, Hoefler J, Bock S, et al: Prostate cancer prevention trial risk calculator 2.0 for the prediction of low- vs high-grade prostate cancer. Urology 2014; **83**: 1362–1367.

24. N. Mottet, R.C.N. van den Bergh, E. Briers, et al: EAU-EANM-ESTRO-ESUR-SIOG Guidelines on Prostate Cancer. 2019. Available at: https://uroweb.org/wp-content/uploads/EAU-EANM-ESUR-ESTRO-SIOG-Guidelines-on-Prostate-Cancer-2019.pdf.

25. Moldovan PC, Van Den Broeck T, Sylvester R, et al: What Is the Negative Predictive Value of Multiparametric Magnetic Resonance Imaging in Excluding Prostate Cancer at Biopsy? A Systematic Review and Meta-analysis from the European Association of Urology Prostate Cancer Guidelines Panel. Eur. Urol. 2017; **72**: 250–266. Available at: http://dx.doi.org/10.1016/j.eururo.2017.02.026, accessed January 6, 2019.

26. Wang RS, Kim EH, Vetter JM, et al: Determination of the Role of Negative Magnetic Resonance Imaging of the Prostate in Clinical Practice: Is Biopsy Still Necessary? Urology 2017; **102**: 190–197. Available at: https://linkinghub.elsevier.com/retrieve/pii/S0090429516308159, accessed April 14, 2019.

27. Elwyn G, Frosch D, Thomson R, et al: Shared Decision Making: A Model for Clinical Practice. J. Gen. Intern. Med. 2012; **27**: 1361–1367. Available at: http://www.ncbi.nlm.nih.gov/pubmed/22618581, accessed April 14, 2019.

28. Gómez-Gómez E, Carrasco-Valiente J, Blanca-Pedregosa A, et al: European Randomized Study

of Screening for Prostate Cancer Risk Calculator: External Validation, Variability, and Clinical Significance. Urology 2017; **102**: 85–91. Available at: http://www.ncbi.nlm.nih.gov/pubmed/27840252, accessed April 13, 2019.

29. Hamoen EHJ, de Rooij M, Witjes JA, et al: Use of the Prostate Imaging Reporting and Data System (PI-RADS) for Prostate Cancer Detection with Multiparametric Magnetic Resonance Imaging: A Diagnostic Meta-analysis. Eur. Urol. 2015; **67**: 1112–1121. Available at: http://www.ncbi.nlm.nih.gov/pubmed/25466942, accessed April 14, 2019.

30. Ankerst DP, Till C, Boeck A, et al: The Impact of Prostate Volume, Number of Biopsy Cores and American Urological Association Symptom Score on the Sensitivity of Cancer Detection Using the Prostate Cancer Prevention Trial Risk Calculator. J. Urol. 2013; **190**: 70–76. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23313212, accessed April 14, 2019.

Tables

Table 1: Clinicopathological characteristics of the cohort

Legends

Figure 1: Area under the receiver operating characteristics curves with corresponding 95%-confidence intervals.

Figure 2: Calibration-in-the-large: Mean predicted amount of csPCa per PCa-RC with bootstrapped 95%-confidence intervals.

Figure 3: Calibration plots with 95%-confidence intervals, intercepts and slopes.

Figure 4: Decision curve analysis for the diagnosis of csPCa.

Table 1 - Clinicopathological characteristics and biopsy results (n=468)

Variable		missing values
Age (years)	64.5 (59 - 68.9)	
PSA (ng/ml)	6.6 (4.5 - 9.9)	
Prostate volume (ml)	48.9 (36 - 63.7)	
Suspicous DRE	77 (16.5)	
Positive family history	50 (10.7)	190 (40.6%)
Previous negative biopsy	145 (31)	
Highest ROI (PIRADS version 2)		
no ROI	87 (18.5)	
2	68 (14.5)	
3	94 (20)	
4	162 (34.6)	
5	57 (12.2)	
Systematic biopsy cores	40 (39-42)	
Targeted biopsy cores per ROI	3 (2-3)	
Biopsy results		
no PCa	220 (47)	
Gleason Score 6	55 (11.8)	
csPCa	193 (41.2)	

Values represented as median with interquartile ranges or as number with percentage of the whole cohort; DRE = digital rectal exam; ROI = region of interest; PCa = prostate cancer; csPCa = clinically significant PCa defined as at least one Gleason pattern ≥4 .